# Active learning of local predictable representations with artificial curiosity

Mathieu Lefort*† and Alexander Gepperth*†
*Computer science and system engineering department - ENSTA ParisTech
828 boulevard des Maréchaux, 91762 Palaiseau Cedex - France
† Inria FLOWERS, Inria Bordeaux Sud-Ouest
200 avenue de la vieille tour, 33405 Talence - France
Email: {mathieu.lefort,alexander.gepperth}@ensta-paristech.fr

*Abstract*—**In this article, we present some preliminary work on integrating an artificial curiosity mechanism in PROPRE, a generic and modular neural architecture, to obtain online, open-ended and active learning of a sensory-motor space, where large areas can be unlearnable. PROPRE consists of the combination of the projection of the input motor flow, using a self-organizing map, with the regression of the sensory output flow from this projection representation, using a linear regression. The main feature of PROPRE is the use of a predictability module that provides an interestingness measure for the current motor stimulus depending on a simple evaluation of the sensory prediction quality. This measure modulates the projection learning so that to favor the representations that predict the output better than a local average. Especially, this leads to the learning of local representations where an input/output relationship is defined [1]. In this article, we propose an artificial curiosity mechanism based on the monitoring of learning progress, as proposed in [2], in the neighborhood of each local representation. Thus, PROPRE simultaneously learns interesting representations of the input flow (depending on their capacities to predict the output) and explores actively this input space where the learning progress is the higher. We illustrate our architecture on the learning of a direct model of an arm whose hand can only be perceived in a restricted visual space. The modulation of the projection learning leads to a better performance and the use of the curiosity mechanism provides quicker learning and even improves the final performance.**

## I. Introduction

Developmental robotics is a recent and active research field that targets the conception of robots that are able to learn to interact with an unknown environment in an autonomous and lifelong open-ended manner, which raises a lot of challenging and yet unsolved problems [3], [4], [5]. The constructivist learning of predictive representations from a sensory-motor data flow is one of them. Especially, this learning does not fit with the classical machine learning framework as large areas of the sensory-motor space are unlearnable such as, for instance, trying to predict the gustatory consequence of moving your arms. One way to deal with learning in such high dimensional sensory-motor spaces is to take inspiration from the infant's development, especially studied in developmental psychology (see [6] e.g.), by providing to the agent intrinsic motivation [7] by mean of an artificial curiosity mechanism. This curiosity mechanism will motivate the agent to explore areas of the sensory-motor space that are interesting for its own development. Various implementations of artificial curiosity were proposed, based on different measures, such as error maximization [8], [9] (areas where the prediction error is the higher are the more interesting) or similarity-based progress maximization [10], [11] (areas where the learning progress is the higher are the more interesting, thus avoiding to get stuck in stochastic areas of the input space) (see [2] for a review).

Especially, IAC [2] (Intelligent Adaptive Curiosity) and its derivative R-IAC [11] (Robust IAC) propose a generic framework for implementing a curiosity mechanism upon a prediction learning model and was applied to the developmental robotics field. This framework consists of the monitoring of the prediction learning progress in various local regions paving the input space. This progress is measured in each region by the difference of accumulated errors in two equal and consecutive sliding temporal windows. In R-IAC, the next input experienced by the prediction learning model is randomly chosen either in the input space (typically 30% of times) or in one of the regions that is picked up with a probability depending on its learning progress. When some predefined number of experiments have been processed in one region, it is split in two along one axis of the input space, so that to maximize the difference between the learning progress in the two newly created regions. For more details on the algorithm, please refer to [11].

PROPRE, that stands for PROjection-PREdiction, is a generic and modular hybrid framework that provides online learning of input data flow representations that are useful to predict another data flow. It combines an unsupervised learning module, based on a self-organizing map (SOM), that autonomously learns representations from the input flow, with a discriminative module that learns the correspondence between the SOM representations and the output by mean of a linear regression. The main originality of PROPRE consists of a predictability module that computed an interestingness measure of the current stimulus, based on the comparison between a prediction quality measure and a sliding threshold. This interestingness measure modulates the generative learning so that to favor the learning of representations that predict the target better than a local average. This predictability modulation mechanism provides a better representations learning, for instance for visual pedestrian pose classification [12]. Moreover, it leads to the gathering of the representations where a relationship with the output is defined, thus avoiding unlearnable areas of the input space [1], so that PROPRE can be well suited for the developmental robotics field.

In this article, we propose to extend the PROPRE capabilities by including a curiosity mechanism in order to simultaneously learn and actively explore an unknown sensory-motor space in a closed perception/action loop. The curiosity mechanism is inspired by the one proposed in R-IAC [11]. For that purpose, each unit of the SOM, that provides a local representation of the input flow based on prototype learning, is associated with a learning progress measure that monitors the model performance evolution in the corresponding Voronoi cell in the input space. Thus, a SOM unit is similar to a region in R-IAC, except that the region is here not defined by a splitting mechanism but dynamically depends on the SOM learning, thus providing a fixed a priori memory usage. The next motor action performed by the system is chosen in some unit prototype neighborhood, the unit being picked up randomly with a probability depending on its learning progress, similarly to what was proposed in R-IAC.

In the next section, we introduce the PROPRE architecture including the new artificial curiosity mechanism. In section III, we show that the curiosity mechanism leads to better asymptotic performance and better performance in almost every time steps when PROPRE learns the direct model of a simple two dimensional planar arm with a limited visual field so that large areas in the motor space are unlearnable. We conclude and discuss possible perspectives of our work in section IV.

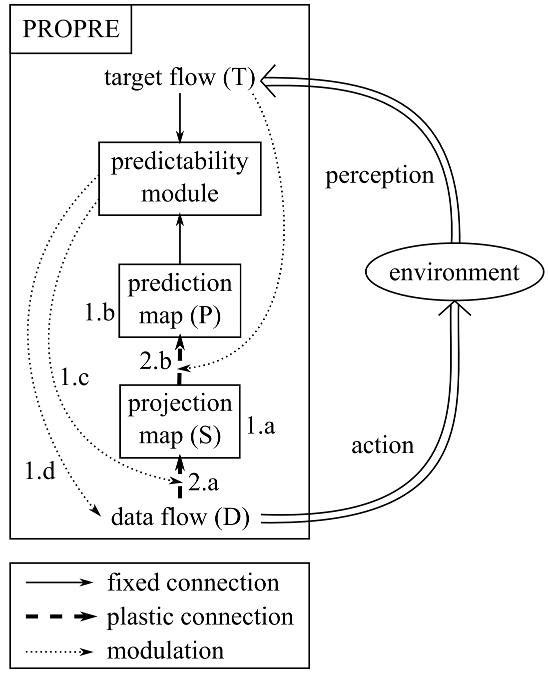


Fig. 1. Data flow processing in PROPRE. The system learns representations of the input data flow ($D$) by combining generative learning ($S$) and discriminative learning of the target ($T$). The predictability module, that monitors the quality of the prediction ($P$), modulates the generative learning so that to favor representations that are better to predict of the input/output relationship than a local average. It also chooses the next input received by the model, depending on the learning progress, closing the perception/action loop.

## II. PROPRE

### A. Architecture

PROPRE is a generic and modular neural paradigm that combines projection and prediction for online learning of input/output relationship from raw data flows. It consists of the interaction between three modules (see figure 1):

- A projection module (see section II-B for details) that transforms the current input stimulus in a low dimensional representation. This module uses the self-organizing map paradigm (SOM) that provides a topological projection of the input space on the map manifold.
- A prediction module (see section II-C for details) that learns to predict the current output target from the representation given by the projection module. This learning is done with a linear regression.
- A predictability module that analysis the quality of the current prediction and plays two roles. First, it modulates the projection module so that to learn representations that are more efficient to predict the target flow than a local average (see section II-D1 for details). Second, it influences the choice of the next input of the model so that to favor areas of the input space leading to learning progress, thus providing an active learning of the input/output relationship (see section II-D2 for details).

Thus, PROPRE is an hybrid architecture that provides online, adaptive, active learning [12] that can also be unsupervised when using PROPRE in a multimodal context where representations learned from each flow try to predict the one of the other flow [13]. Besides, depending on the way to use the SOM activity to define the representations, PROPRE can be seen as an extension of a radial basis function network [14].

From a computational point of view, the next input stimulus of the model is the (proprioception of the) motor command that is actively determined depending on the predictability module (equation 1.d in figure 1). The value of the target corresponds to the perception of the environment resulting from the execution of this motor command. The input stimulus is processed in the model by a feed-forward evaluation of each module activity (equations 1.a-b-c). Then, the plastic connections are updated depending on the corresponding modules activities (equations 2.a-b) and so on.

### B. Projection

In our previous articles [1], [12], the projection step consisted of the classical self-organizing map model proposed by Kohonen [15]. This model is related to the minimization of the quantization error - plus a topological term - so that the distribution of the prototypes tends to be similar to the one of the input data [15], [16]. In this article, we introduce a curiosity mechanism to actively choose the next input (see section II-D2 for more details), thus deeply modifying the input data distribution. The coupling of this curiosity mechanism with the Kohonen learning rule creates an undesirable dynamic attractor corresponding to the convergence of all prototypes in a very tiny area of the input space. Indeed, the input are chosen around some of the learned prototypes, i.e. in a subspace of the one mapped by the SOM, leading to the mapping of this subspace by the SOM and so on and so forth.

In order to tackle this problem, in this article, we use a Dynamic Self-Organizing Map (DSOM) [17] for the prediction module. DSOM model, compared to the one of Kohonen, includes the distance between the stimulus and the best matching prototype in the computation of the learning rate and the Gaussian neighborhood (see equation 2.a) so that the model is only slightly changed if a prototype is already close to the stimulus. Thus, even if the curiosity mechanism concentrates the input in some new interesting subspace of the input space for some time, it will have few consequences on the previously learned projection as DSOM does not reduce a quantization error but tends to map the input space [17]. Thus, DSOM provides a kind of incremental learning, dealing with a fixed number of units in the map, necessary to the use of a curiosity mechanism [11].

In practice, $S$ is a discrete bi-dimensional square grid of neurons, each one receiving the input data flow $D$ (see figure 1). Each prototype of a unit $x$ in the map, denoted $\mathbf{w_{SD}}(x,t)$, is updated at each time step $t$ with the following equation, corresponding to the one of DSOM modulated with the interestingness measure $I(t)$ provided by the predictability module (see section II-D1):

$$\Delta \mathbf{w_{SD}}(x,t) = d^*(t)\eta I(t) e^{\frac{-||x-x^*(t)||_2^2}{2(d^*(t)\sigma)^2}} (\mathbf{D}(t) - \mathbf{w_{SD}}(x,t)) \quad (2.a)$$

with $\eta$ the learning rate and $\sigma$ the variance of the Gaussian neighborhood radius which are both constant to provide a lifelong learning. $||\cdot||_2$ is an euclidean distance in the map, $x^*(t)$ is the best matching unit defined as the unit whose prototype is the closest to the current input $\mathbf{D}(t)$, i.e. $x^*(t) = \arg\min_x |\mathbf{w_{SD}}(x,t) - \mathbf{D}(t)|_2$ and $d^*(t)$ is the distance between this best matching unit and the current input, i.e. $d^*(t) = |\mathbf{w_{SD}}(x^*(t),t) - \mathbf{D}(t)|_2$, with $|\cdot|_2$ an euclidean distance on the input space.

The activity of any unit $x$ of the map is defined as:

$$S(x,t) = \begin{cases} 1 \text{ if } x = x^*(t) \\ 0 \text{ otherwise} \end{cases} \quad (1.a)$$

This classical activation function [14] provides a prediction that only depends on the best matching unit (see section II-C). Thus, it is consistent with the monitoring of the prediction quality, computed in each Voronoi cell of DSOM units, processed in the predictability module (see section II-D).

### *C. Prediction*

The projection activity $\mathbf{S}(t)$ is used to compute a prediction $\mathbf{P}(t)$ of the target data flow $\mathbf{T}(t)$ at time $t$. The activity of a unit $x$ in $P$ is computed as a weighted sum of the $S$ activity:

$$P(x,t) = \sum_y w_{PS}(x,y,t) S(y,t) \quad (1.b)$$

with $w_{PS}(x,y,t)$ the weight from the units $y$ in $S$ to $x$ in $P$.

The connection weights between $S$ and $P$ are learned with a classical stochastic gradient descent implementation of a linear regression [18], which minimizes the mean square error between the prediction $\mathbf{P}(t)$ and the current target value $\mathbf{T}(t)$. Thus, the weights are updated with the following equation:

$$\Delta w_{PS}(x,y,t) = \eta' S(y,t)(T(x,t) - P(x,t)) \quad (2.b)$$

with $\eta'$ the constant learning rate.

### *D. Predictability module*

The predictability module monitors the quality $Q(t)$ of the current prediction $\mathbf{P}(t)$ with respect to the true target value $\mathbf{T}(t)$ with a simple and generic measure:

$$Q(t) = \frac{P(z^*,t)}{\sum_z P(z,t)} \text{ with } z^* = \arg\max_z \; T(z,t)$$

In our experiments, the target value represents the visual position of the end effector in a matrix of pixels, when the hand is visible (see section III). Thus, $Q(t)$ represents the percentage of prediction of the right position of the hand. This measure was also used for classification learning [12] or multimodal regression [13].

*1) Projection learning modulation:* In each Voronoi cell[1] associated to a unit $x$ of DSOM, we compute online the average prediction quality $\theta(x,t)$ over a sliding window:

$$\theta(x,t) = \begin{cases} (1-\tau)\theta(x,t-1) + \tau Q(t) \text{ if } x = x^*(t) \\ \theta(x,t-1) \text{ otherwise} \end{cases} \quad (1)$$

with $\tau$ the smoothing factor and $x^*(t)$ the best matching unit at time $t$ (see section II-B).

The interestingness measure of the current stimulus for predicting the target, that modulates the projection learning (equation 2.a), is defined as:

$$I(t) = \begin{cases} Q(t) - \theta(x^*(t),t) \text{ if } Q(t) >= \theta(x^*(t),t) \\ 0 \text{ otherwise} \end{cases} \quad (1.c)$$

Thus, the current stimulus is only learned if it provides a prediction locally more accurate than the average and the interestingness measure depends directly on this accuracy difference to the local average.

*2) Artificial curiosity:* As proposed in [2], [11], the artificial curiosity mechanism is based on the monitoring of learning progress of the model on various local areas paving the input space, here the Voronoi cells associated with the DSOM units. In practice, the learning progress $LP(x,t)$ of each unit $x$ of the DSOM at time $t$ is computed as the difference between two average prediction quality computed online with different smoothing factors $\tau$ and $\tau'$ ($\tau' < \tau$). The learning progress itself is averaged over time with $\tau_{LP}$ as smoothing factor:

[1] In practice, as the system continuously learns, unit prototypes are continuously smoothly modified so that $\theta(x,t)$ is only a reasonable approximation of the average prediction quality in the Voronoi cell of the unit.

$$LP(x,t) = \begin{cases} (1-\tau_{LP})LP(x,t-1) + \tau_{LP}(\theta(x,t) - \theta'(x,t)) \\ \qquad\qquad\qquad\qquad\qquad \text{if } x = x^*(t) \\ LP(x,t-1) \text{ otherwise} \end{cases}$$

$$\theta'(x,t) = \begin{cases} (1-\tau')\theta'(x,t-1) + \tau' Q(t) \text{ if } x = x^*(t) \\ \theta'(x,t-1) \text{ otherwise} \end{cases}$$

The average prediction quality $\theta(x,t)$, computed on the shortest time window, is the one used for the interestingness measure (see previous section).

The next action to perform is uniformly chosen in the input space with some probability (25% in our experiments) for exploration, otherwise it is actively determined by the curiosity mechanism. For that purpose, the model computes for each unit $x$ of the DSOM a probability $p(x,t)$, depending of its learning progress, to be picked up :

$$p(x,t) = \frac{|LP(x,t)|}{\sum_{x'} |LP(x',t)|}$$

Let denote $\hat{x}(t)$ the effectively randomly picked up unit by the model at time $t$. Then, the next input $\mathbf{D}(t+1) = (D(i,t+1))_i$ of the model is obtained by adding to each dimension of the prototype $\mathbf{w_{SD}}(\hat{x}(t),t) = (w_{SD}(\hat{x}(t),i,t))_i$, associated with the randomly chosen unit $\hat{x}(t)$, a value in $[-r,+r]$ picked up with an uniform probability:

$$D(i,t+1) = w_{SD}(\hat{x}(t),i,t) + \sigma_i \qquad (1.d)$$

with $\sigma_i \sim \mathcal{U}(-r,r)$.

It has to be noticed that the interestingness measure (section II-D) can be interpreted as the rectified temporal evolution of the prediction quality measure between short term (current quality $Q(t)$) and medium term ($\theta(x,t)$) averages. Thus, both the projection modulation and the artificial curiosity are based on the same mechanism operating at different time scales and their intertwining leads to the emergent properties of PROPRE.

## III. Experiments and Results

### A. Protocol

We tested our PROPRE architecture on the learning of the direct model of a simple simulated robotic planar arm (see figure 2). In our setup, the hand is seen by a $5 \times 10$ pixels matrix covering a restricted part of the reachable area so that the direct model is learnable only for around 20% of the motor space (see figure 3). This experiment, a simplified and adapted version of the hand-eye-clouds experiment proposed in [11], illustrates one of the main problem in developmental robotic: exploration of a sensory-motor space where large areas are unlearnable.

With the curiosity mechanism, the next motor command is chosen in some neighborhood of a learned prototype (see section II-D2) so that it can be outside of the joint limits. In this case, the motor command is bounded within the joint limits and the input/output couple provided to PROPRE is the proprioception and visual perception of the effectively performed motor command.

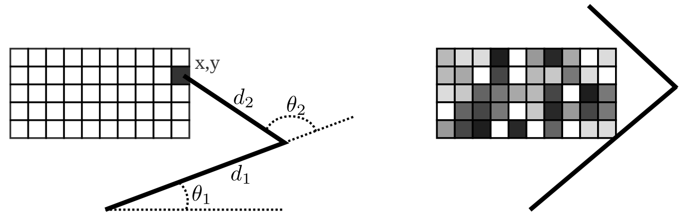


Fig. 2. A robot moves its arm ($d_1 = 3$ and $d_2 = 2.5$, the units are arbitrary) in a plan depending on an input joint motor command ($(\theta_1, \theta_2) \in [0, \pi[^2$). It can only see its hand with a pixel matrix providing a restricted visual field. If it sees its arm (left), the corresponding pixel is set to 1, otherwise (right) the visual perception is a white noise with amplitude 1.

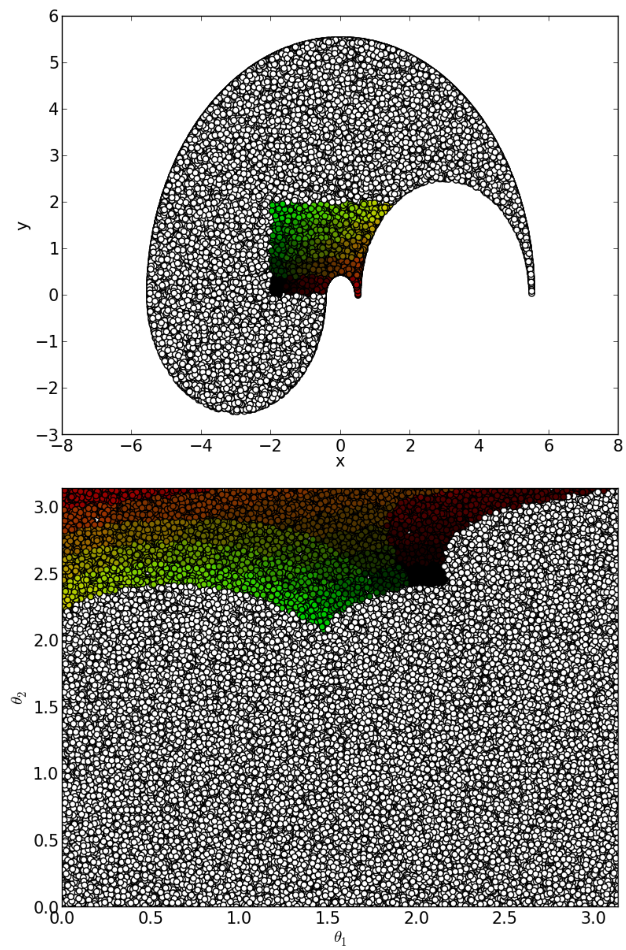


Fig. 3. Example of visual (top) and motor (down) couples. White area correspond to an invisible hand, whereas each colored area corresponds to some pixel activated in the visual matrix (better viewed in color).

### B. Results

In order to study the influence of the predictability module on the performance, we tested three different architectures:

- no modulation of the projection learning (i.e. $\forall t,\ I(t) = 1$) and no curiosity, denoted DSOM+LR,
- modulation of the projection learning and no curiosity, denoted PROPRE,
- modulation of the projection learning and curiosity, denoted PROPRE ACTIVE.

When no curiosity mechanism is used, the motor actions are randomly chosen in the input space with an uniform distribution. When used, the curiosity mechanism chooses 75% of the performed actions, the others are randomly picked up in the motor space.

For each architecture, the DSOM size was $10 \times 10$, the prediction learning rate was set to $\eta' = 10^{-3}$ (equation 2.b). We independently tuned, with reasonable effort, the projection parameters for DSOM+LR ($\sigma = 3$, $\eta = 1$) and PROPRE ($\sigma = 4$, $\eta = 3$) (equation 2.a). Then, we tune the smoothing factor used in the interestingness measure, $\tau = 10^{-3}$ (equation 1.c). These parameters, found for PROPRE, are also used for PROPRE ACTIVE is order to fairly quantify the influence of the curiosity mechanism. Finally, we tuned the curiosity mechanism and set the long term smoothing factor $\tau' = 10^{-4}$ and the range $r = 0.35$ (see section II-D2). Moreover, the initial weights of each unit of DSOM were initialized in $[(\pi-1)/2, (\pi+1)/2]^2$, i.e. in the center of the motor space but outside of the motor area providing a visible hand.

In order to evaluate the performance of the models, we defined a benchmark of 10000 motor/visual couples where the hand is visible and recorded an error each time the maximum of the prediction did not correspond to the real position of the hand. The results are presented in figure 4. By the way, similar qualitative results are obtained with the same parameters when putting a random pixel at 1 and the others at 0, instead of white noise, when the hand is not visible during learning.

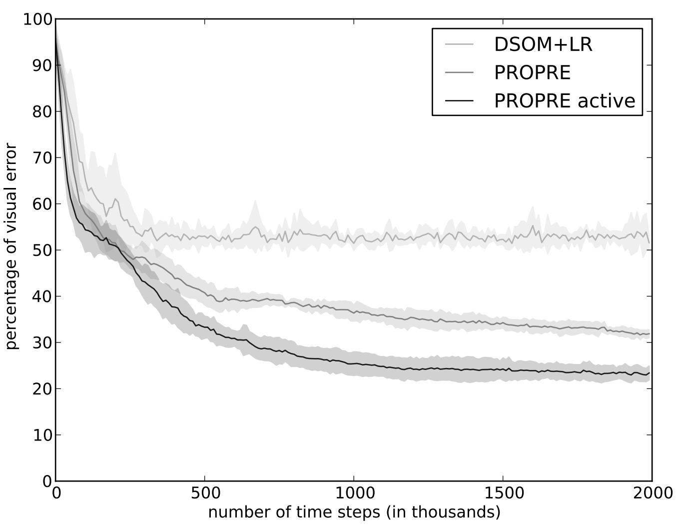


Fig. 4. Average and variance (over 10 simulations) of the temporal evolution of performance obtained by the different models.

We can clearly see that the modulation of the projection learning provides better performance at each time step compared to the DSOM+LR system. This confirms the results we obtain in [1] but when using a Kohonen SOM module for the projection. More interestingly, we can observe that the curiosity mechanism, other things being equal, has also a significant influence on the performance. Indeed, the PROPRE ACTIVE model is the one that has better performance at almost every time step including the better asymptotic performance. This improvement is obtained because the curiosity mechanism performs mainly motor actions providing a visible hand, 60% in average over the ten simulations[2], whereas only 20% of the total motor space leads to a visible hand.

On figure 5, we illustrate an example of the temporal evolution of the repartition of the performed motor actions obtained when using the curiosity mechanism. We can observe that the motor actions seem to first mainly concentrate on the orange-red areas of figure 3 before spreading over all the motor area leading to a visible hand. These orange-red areas are the easiest ones to learn as the activation of one specific pixel can be obtained by a larger range of motor actions (compare the relative size of the colored areas in the motor space in figure 3). Thus, this seems to indicate that the curiosity mechanism leads to a developmental trajectory of the model from simple to more complicated learning as already found when using the IAC framework [11]. A more extensive study, including statistical results over various trials, is necessary to confirm this developmental trajectory in PROPRE.

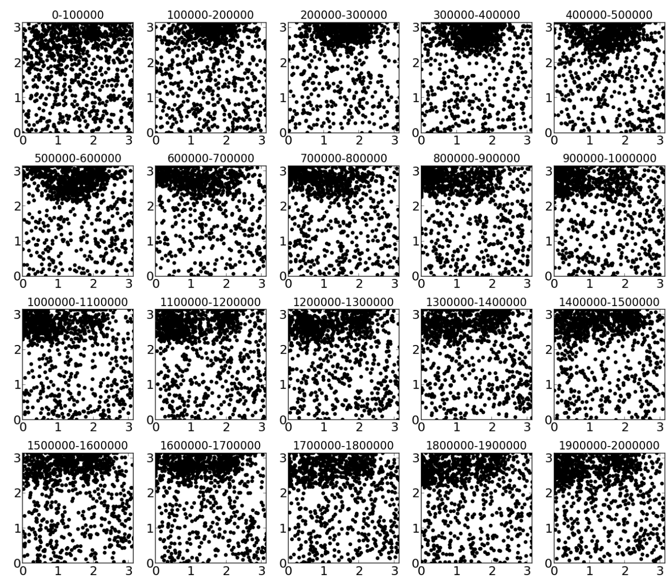


Fig. 5. From top left to bottom right: temporal evolution of the distribution of motor actions executed by the system using the curiosity mechanism in one of the simulations. Each figure represents, in the motor space, the 100000 actions executed by the system during the time window mentioned above the correspond figure.

## IV. CONCLUSION AND PERSPECTIVES

PROPRE is a neural paradigm for online and open-ended learning of input/output relationship from raw data flows. It combines a generative learning, by projecting the input space on a self-organizing map - here the DSOM model, with a discriminative learning, by mean of a linear regression

[2]For exploration purpose, 25% of actions are randomly chosen in the motor space (see section II-D2), leading to around 20% of total actions performed by the system in the unlearnable area. So that half of the actions leading to an invisible hand are due to the exploration mechanism.

of the output. The main originality of PROPRE is the use of a predictability module that monitors the quality of the prediction, with a simple measure, and modulates the projection learning so that to favor representations that predict the output better than a local average. Especially, when the input/output relationship in only defined in some areas of the input space, this predictability modulation leads to the gathering of representations in these learnable areas [1], an important property for developmental robotics.

In this article, we propose to integrate in PROPRE an artificial curiosity mechanism, derived from the IAC paradigm [2], based on the monitoring of learning progress in various regions paving the input space. For that purpose, we compute the temporal evolution of the prediction quality obtained in each Voronoi cell of the DSOM by the difference between two averages computed online with different smoothing factors. The next input is randomly chosen, with an uniform distribution, in some neighborhood of the prototype of one DSOM unit picked up with a probability depending of its associated learning progress. Thus, PROPRE simultaneously learns to represent efficiently the input space and actively explores around learned areas, where it seems the more promising (through the curiosity mechanism monitoring the learning progress), in a closed perception/action loop.

We tested this architecture on the learning of a direct model of a two degrees of freedom planar arm, whose hand position is perceived by a matrix of pixel only covering a limited subspace of the reachable space. The modulation of the projection learning by the predictability module provides a better visual prediction performance. The additional use of the curiosity mechanism provides an even better asymptotic performance. Moreover, by performing mainly motor commands providing a visible hand, the curiosity mechanism reduces the number of actions needed to obtain a defined performance.

These very promising results on a simple setup open the way to test PROPRE for the learning and exploration of more realistic sensory-motor spaces, especially high-dimensional and redundant. We also want to improve the exploration part of the curiosity mechanism, which require yet to know a priori the boundaries of the motor space, for example using social guidance. Moreover, IAC was improved by two paradigms including some features whose integration in PROPRE can lead to interesting perspectives.
First, R-IAC (Robust IAC) [11] monitors the learning progress at several scales by using a tree structure of regions. One way to include this feature in PROPRE can be to use a tree of self-organizing maps, a structure that was already studied to obtain different granularities in the mapping (see [19] e.g.). This question can be coupled with the study of the differences between having an increasing number of fixed regions, as in IAC and its derivatives, versus a fixed number of dynamic regions, as in PROPRE ACTIVE.
Second, SAGG-RIAC (Self-Adaptive Goal Generation RIAC) [20] monitors a competence progress, i.e. the ability to reach some goal in the output space, instead of a learning progress, i.e. the ability to predict the consequence of a motor action. This competence-based curiosity seems to be particularly efficient in highly redundant sensory-motor spaces. PROPRE can be use for multimodal learning, by providing as output of a data flow processing the representations learned from another data flow [1]. Thus, one way to integrate competence-based curiosity in PROPRE could be to learn representations from the sensory and the motor flows and to associate the curiosity mechanism to the sensory representations instead of the motor representations as in this article. Moreover, both mechanisms could be coupled and alternatively used depending on the sensory-motor area currently explored.

## REFERENCES


[1] M. Lefort and A. Gepperth, "Learning of local predictable representations in partially structured environments." in *International Joint Conference on Neural Networks*, 2015.

[2] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *Evolutionary Computation, IEEE Transactions on*, vol. 11, no. 2, pp. 265–286, 2007.

[3] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: a survey," *Connection Science*, vol. 15, no. 4, pp. 151–190, 2003.

[4] F. Guerin, "Learning like a baby: a survey of artificial intelligence approaches," *The Knowledge Engineering Review*, vol. 26, no. 02, pp. 209–236, 2011.

[5] J. Weng, "Developmental robotics: Theory and experiments," *International Journal of Humanoid Robotics*, vol. 1, no. 02, pp. 199–236, 2004.

[6] J. Piaget, M. Cook, and W. Norton, *The origins of intelligence in children*. International Universities Press New York, 1952, vol. 8, no. 5.

[7] R. M. Ryan and E. L. Deci, "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemporary educational psychology*, vol. 25, no. 1, pp. 54–67, 2000.

[8] A. G. Barto, S. Singh, and N. Chentanez, "Intrinsically motivated learning of hierarchical collections of skills," in *Proc. 3rd Int. Conf. Development Learn*, 2004, pp. 112–119.

[9] S. Thrun, "Exploration in active learning," *Handbook of Brain Science and Neural Networks*, pp. 381–384, 1995.

[10] J. Schmidhuber, "Curious model-building control systems," in *Neural Networks, 1991. 1991 IEEE International Joint Conference on*. IEEE, 1991, pp. 1458–1463.

[11] A. Baranès and P.-Y. Oudeyer, "R-iac: Robust intrinsically motivated exploration and active learning," *Autonomous Mental Development, IEEE Transactions on*, vol. 1, no. 3, pp. 155–169, 2009.

[12] M. Lefort and A. Gepperth, "Propre: Projection and prediction for multimodal correlations learning. an application to pedestrians visual data discrimination." in *International Joint Conference on Neural Networks*. IEEE, 2014.

[13] M. Lefort, T. Kopinski, and A. Gepperth, "Multimodal space representation driven by self-evaluation of predictability," in *ICDL-EPIROB-The fourth joint IEEE International Conference on Development and Learning and on Epigenetic Robotics*, 2014.

[14] T. Hecht, M. Lefort, and A. Gepperth, "Using self-organizing maps for regression: the importance of the output function."

[15] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[16] M. Cottrell, J. Fort, and G. Pagès, "Theoretical aspects of the som algorithm," *Neurocomputing*, vol. 21, no. 1-3, pp. 119–138, 1998.

[17] N. Rougier and Y. Boniface, "Dynamic self-organising map," *Neurocomputing*, vol. 74, no. 11, pp. 1840–1847, 2011.

[18] C. Bishop and N. Nasrabadi, *Pattern recognition and machine learning*. springer New York, 2006, vol. 1.

[19] C. A. Astudillo and B. J. Oommen, "Topology-oriented self-organizing maps: a survey," *Pattern Analysis and Applications*, vol. 17, no. 2, pp. 223–248, 2014.

[20] A. Baranes and P.-Y. Oudeyer, "Active learning of inverse models with intrinsically motivated goal exploration in robots," *Robotics and Autonomous Systems*, vol. 61, no. 1, pp. 49–73, 2013.